DTU Library

Modeling comodulation masking release using an equalization-cancellation mechanism

Piechowiak, Tobias; Ewert, Stephan D.; Dau, Torsten

Published in:
Acoustical Society of America. Journal

Link to article, DOI:
10.1121/1.2534227

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Modeling comodulation masking release using an equalization-cancellation mechanism

Tobias Piechowiak, Stephan D. Ewert,[a] and Torsten Dau[b]
Centre for Applied Hearing Research, Acoustic Technology, Ørsted•DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

(Received 30 August 2006; revised 4 January 2007; accepted 5 January 2007)

This study presents an auditory processing model that accounts for the perceptual phenomenon of comodulation masking release (CMR). The model includes an equalization-cancellation (EC) stage for the processing of activity across the audio-frequency axis. The EC process across frequency takes place at the output of a modulation filterbank assumed for each audio-frequency channel. The model was evaluated in three experimental conditions: (i) CMR with four widely spaced flanking bands in order to study pure across-channel processing, (ii) CMR with one flanking band varying in frequency in order to study the transition between conditions dominated by within-channel processing and those dominated by across-channel processing, and (iii) CMR obtained in the "classical" band-widening paradigm in order to study the role of across-channel processing in a condition which always includes within-channel processing. The simulations support the hypothesis that within-channel contributions to CMR can be as large as 15 dB. The across-channel process is robust but small (about 2–4 dB) and only observable at small masker bandwidths. Overall, the proposed model might provide an interesting framework for the analysis of fluctuating sounds in the auditory system. [DOI: 10.1121/1.2534227]

PACS number(s): 43.66.Ba, 43.66.Mk [JHG] Pages: 2111–2126

I. INTRODUCTION

Many properties of auditory masking can be understood in terms of the responses of the basilar membrane within the inner ear. Each part of this membrane behaves like a filter that responds to a limited range of frequencies. When trying to detect a sinusoidal tone in background noise, it has been proposed that listeners use the output of a single auditory filter tuned to the frequency of the tone (Fletcher, 1940). That filter passes the tone at full intensity, but rejects most of the background noise. Although this theory can account for many aspects of masking, Hall *et al.* (1984a) and others showed that, when comodulated maskers were used, some of the results can be explained only if it is assumed that stimulus information is processed across the outputs of auditory filters. In fact, humans are often much better at detecting signals in comodulated maskers than in white noise, an effect called comodulation masking release (CMR; Hall *et al.*, 1984a). Various experiments on CMR have demonstrated that the human auditory system can exploit coherent envelope fluctuations very effectively and that substantial reductions in signal threshold can be the result. Since coherent across-frequency modulation is common in speech, music, animal vocalization and environmental noise, the ability to process such information is thought to be a powerful survival strategy in the natural world which aids in the detection of target sounds in the presence of competing sounds.

CMR was demonstrated initially by Hall *et al.* (1984a). In their "band-widening" experiment, the detection of a tone was measured as a function of the bandwidth of a noise masker, keeping the spectrum level constant. They used two types of maskers. One was a random noise with irregular fluctuations in amplitude that are independent in different frequency regions. The other was a random noise which was amplitude modulated using a low-pass filtered noise as a modulator. This modulation resulted in slow fluctuations in the amplitude of the noise that were the same in different frequency regions. For the random noise, the signal threshold increased as the masker bandwidth increased up to about the critical bandwidth at that frequency and then remained constant, as expected from the classical power spectrum model of masking (Fletcher, 1940; Patterson and Moore, 1986). The pattern for the modulated noise was quite different. Here, the threshold decreased as the bandwidth was increased beyond about 100 Hz (for a signal frequency of 2 kHz); thus, adding more noise to the masker made the signal easier to detect. This suggested that subjects may compare the outputs of different auditory filters to enhance signal detection. The fact that the decrease in threshold with increasing bandwidth only occurred with the modulated noise indicated that fluctuations in the masker are critical and that the fluctuations need to be correlated across frequency bands.

In a second class of experiments, CMR was demonstrated by using narrow bands of noise (of typically 20–50 Hz width), which inherently have relatively slow amplitude fluctuations. One band, the on-frequency band, was centered at the signal frequency. A second band, the flanker band, was placed remote from the signal frequency. When the flanking band was uncorrelated with the on-frequency band, there was typically no effect on signal threshold. How-

[a]Current address: Carl von Ossietzky Universität Oldenburg, Medizinische Physik D-26111 Oldenburg, Germany
[b]Author to whom correspondence should be addressed. Electronic mail: tda@oersted.dtu.dk

ever, when the flanking band was correlated with the on-frequency band, a flanking band produced a release from masking (Hall *et al.*, 1984a; Schooneveldt and Moore, 1987; Cohen and Schubert, 1987). CMR was also found even if the signal and on-frequency band were presented to one ear and the flanking band to the other ear (Schooneveldt and Moore, 1987; Cohen and Schubert, 1987).

Even though CMR has been investigated in a number of studies, the underlying mechanisms are still not clear. It has generally been assumed that CMR results from across-channel comparisons of temporal envelopes. Alternatively, it has been suggested that analysis of the output of a broad initial predetection filter, which encompasses frequencies generally thought to fall into separate auditory filters, can account for certain aspects of CMR (Berg, 1996). However, Buss *et al.* (1998) and Buss and Hall (1998) provided evidence against such a broad predetection filter; their results were, instead, consistent with an initial stage of auditory (bandpass) filtering. Other studies have proposed that within-channel cues, i.e., information from only the one auditory channel tuned to the signal frequency, can account for a considerable part of the effect in some conditions, which means that within-channel processing can lead to an overestimation of "true" across-channel CMR (e.g., Schooneveldt and Moore, 1987). This was supported by simulations of data from the band-widening experiment, using a modulation filterbank analysis of the stimuli at the output of the auditory filter tuned to the signal frequency (Verhey *et al.*, 1999). Additionally, for the CMR experiments using flanking bands, McFadden (1986) pointed out that it is imprecise to assume that one channel is receiving only the on-frequency band plus signal and another channel is receiving only the flanking band. Often, the two bands will be incompletely resolved. When this happens, the resulting waveform may contain envelope fluctuations resulting from beats between the carrier frequencies of the on-frequency and the flanker bands. These beats can facilitate signal detection without across-channel comparisons being involved. Thus, at least part of the masking release can be explained in terms of the use of within-channel rather than across-channel cues. Taken together, across-channel CMR appears to be a robust, but relatively small effect, which was found in monotic and dichotic conditions.

A recent study on effects of auditory grouping on CMR (Dau *et al.*, 2005) supported two forms of CMR. In their study, the effects of introducing a gating asynchrony between on-frequency and flanker bands or a stream of preceding (precursor) or following (postcursor) flanker bands were studied for conditions of CMR. Using widely (one octave) spaced flanking bands, CMR effects were eliminated by introducing a gating asynchrony and by introducing the pre-or postcursor flanking bands. Using narrowly spaced flanking bands (one-sixth octave), CMR was not affected by any of the stimulus manipulations. Their results supported the hypothesis that one form of CMR is based on within-channel mechanisms (Schooneveldt and Moore, 1987; Verhey *et al.* (1999)), determined by the envelope statistics. The fact that this effect was not susceptible to manipulations by auditory grouping constraints is in line with the assumption that the mechanism is peripheral in nature, based on the physical interaction of stimulus components within an auditory channel. The other form of CMR, mainly based on "true" across-channel comparisons, appeared to be dependent on auditory grouping constraints, consistent with the results from Grose and Hall (1993).

Several hypotheses have been suggested about the nature of the across-channel mechanism underlying CMR. One hypothesis is based on the assumption that the addition of the signal to the on-frequency masker band leads to a change in the modulation depth in the auditory filter centered at the signal frequency. By comparing this modulation depth to that of other auditory filters for which the modulation depth is unaltered, subjects would increase their sensitivity to the presence of the signal (Hall, 1986). A different explanation for CMR was proposed by Buus (1985), who suggested that the comodulated flanker band(s) provide valuable information about the moments at which the masker band has a relatively low energy. By attributing more weights to these valleys in the masker, the effective signal-to-noise ratio increases and detection improves. This mechanism was called "listening in the valleys." Also proposed by Buus (1985) was an equalization-cancellation (EC) mechanism, originally introduced by Durlach (1963), to account for various binaural masking release data. According to this mechanism, the envelope of the masker and flanking band are first equalized and then subtracted. The output of such a mechanism might have a considerable increase in the signal-to-noise ratio provided that the masker and the flanking bands are comodulated.

A fourth mechanism has been proposed by Richards (1987), where it was assumed that the cross *covariance* between the envelopes of the masker and the flanking bands is used for signal detection. The envelope cross covariance decreases when adding a signal to the masker and this cue might be used by the human auditory system. However, this model was later rejected because it was not compatible with experimental data by Edins and Wright (1984). They used two 100% sinusoidally amplitude modulated sinusoids of different frequencies, and the subjects had to detect the in-phase addition of a sinusoid to one of the SAM sinusoids. The cross covariance is not changed even though the modulation pattern is altered by the addition of the sinusoid. Thus, if changes in the cross covariance were essential for receiving CMR, this type of stimulus should not lead to a CMR. This, however, was in contrast to their data, which clearly showed CMR.

Later, van de Par and Kohlrausch (1998b) and van de Par (1998) found that CMR can better be described in terms of an envelope cross *correlation* mechanism than an envelope cross covariance mechanism. Their study was motivated by earlier findings by Bernstein and Trahiotis (1996) which showed that cross correlation was more successful than cross covariance when studying binaural detection phenomena. At high frequencies where these experiments were carried out, similar mechanisms may indeed underly monaural CMR and binaural masking level differences (BMLD, van de Par and Kohlrausch, 1998a). Moreover, the EC mechanism, which has been used to account for BMLD, was shown to be equivalent to a decision mechanism based on cross correlation (Domnitz and Colburn, 1977; Green, 1992).

While potential mechanisms underlying CMR have been discussed in several studies, predictions that quantify the (relative) contributions of across- versus within-channel processing in different types of experiments have not been provided. The purpose of the present study was therefore to develop and evaluate a model that accounts for both effects in CMR. The modulation filterbank model by Dau *et al.* (1997a, b) was considered as the modeling framework. This model was used earlier to analyze within-channel cues in CMR obtained in the band-widening experiment (Verhey *et al.*, 1999), and applied to a variety of other detection and masking conditions, including tone-in-noise detection, modulation detection, and forward masking. In the Verhey *et al.* (1999) study, the model was exclusively tested in the band-widening experiment of CMR. The results from the simulations, performed only in the auditory channel tuned to the signal frequency, suggested that essentially no across-channel processing is involved in this type of CMR condition. Instead, temporal within-channel cues such as beating between components, evaluated by the modulation filterbank model, appear to account for the masking release in the model simulations. However, since the model does not contain any explicit across-channel processing, it will not be able to account for any true across-channel CMR. In the present study, an EC-based circuit was integrated into an extended version of the modulation filterbank model whereby the EC processing was assumed to take place at the level of the internal representation of the stimuli *after* modulation filtering.

First, the structure of the across-channel modulation filterbank model is described. The model is then evaluated in several experimental conditions: (i) CMR with four widely spaced flanker bands to study pure across-channel CMR (Experiment 1), (ii) CMR with one flanking band varying in frequency in order to study the transition between conditions dominated by within-channel processing and those dominated by across-channel processing (Experiment 2), and (iii) CMR obtained in the band-widening paradigm in order to study the contribution of across-channel processing in a condition which always includes within-channel processing (Experiment 3). For direct comparison, experimental data were obtained in the same conditions with exactly the same stimuli and using exactly the same threshold algorithm as in the simulations. The results and implications for further modeling work are discussed.

II. MODEL

The model presented here is based on the monaural detection model of Dau *et al.* (1997a). The original model was designed to account for signal detection data in various psychoacoustic conditions. It has proven successful in predicting data from spectral and spectro-temporal masking (Verhey *et al.*, 1999; Derleth and Dau, 2000; Verhey, 2002), nonsimultaneous masking (Dau *et al.*, 1996, 1997a; Derleth, *et al.*, 2001) and modulation detection and masking (Dau *et al.*1997a, b, 1999; Ewert and Dau, 2004). In the meantime, an additional model of amplitude modulation (AM) processing, the envelope power spectrum model (EPSM) has been developed (Ewert and Dau, 2000; Ewert *et al.*, 2002), based on Viemeister (1979) and Dau *et al.* (1999). The EPSM has a much simpler structure than the abovementioned processing model. It is similar to Viemeister's (1979) leaky-integrator model but assumes modulation bandpass filters instead of a single modulation lowpass filter. It consists of only three stages: Hilbert-envelope extraction, modulation bandpass filtering, and a decision stage based on the long-term, mean integrated envelope power. This model does not include any effects of peripheral filtering and adaptation, and timing information (as reflected in the envelope phase and modulation beatings) is neglected. While the EPSM demonstrated in a straightforward and intuitive way the need for modulation-frequency selective processing and can account for modulation masking data, it is conceptually less general than the perception model (Dau *et al.*, 1996, 1997a).

The model as described in Dau *et al.* (1997a), which forms the basis for the model developed here, consists of the following steps: Peripheral filtering, envelope extraction, nonlinear adaptation, modulation filtering, and an optimal detector as the decision device. To simulate the bandpass characteristic of the basilar membrane, the gammatone filterbank (Patterson *et al.*, 1987) is used. At the output of each peripheral filter, the model includes half-wave rectification and low-pass filtering at 1 kHz. While the fine structure is preserved for low frequencies, for high center frequencies this stage essentially preserves the envelope of the signal. Effects of adaptation are simulated by a nonlinear adaptation circuit (Püschel, 1988; Dau *et al.*, 1996). For a stationary input stimulus, this stage creates a compression close to logarithmic. With regard to the transformation of envelope fluctuations, the adaptation stage transforms the AM depth of input fluctuations with rates higher than about 2 Hz almost linearly. The stimuli at the output of the adaptation stage for each channel are then processed by a linear modulation filterbank. The lowest modulation filter is a second-order butterworth lowpass filter with a cutoff frequency of 2.5 Hz. For frequencies above 5 Hz there is an array of bandpass filters with a quality factor of $Q=2$. Modulation filters with a center frequency above 10 Hz only output the Hilbert envelope of the modulation filters, introducing a nonlinearity into the modulation processing through which the phase of the envelope is not preserved for the filters above 10 Hz. To model a limit of resolution, an internal noise with a constant variance is added to the output of each modulation filter. In the decision process, a stored, normalized temporal representation of the signal to be detected (the template) is compared with the actual activity pattern by calculating the cross correlation between the two temporal patterns (Dau *et al.*, 1996, 1997a). This is comparable to a "matched filtering" process (Green and Swets, 1966).

For the processing of arbitrary input stimuli, the function of the model can be considered as being separated in two (parallel) paths: (i) The stimulus representation after nonlinear adaptation is low-pass filtered at a cutoff frequency of 2.5 Hz, thereby essentially extracting the stimulus energy. With this processing alone, the model would be acting simi-

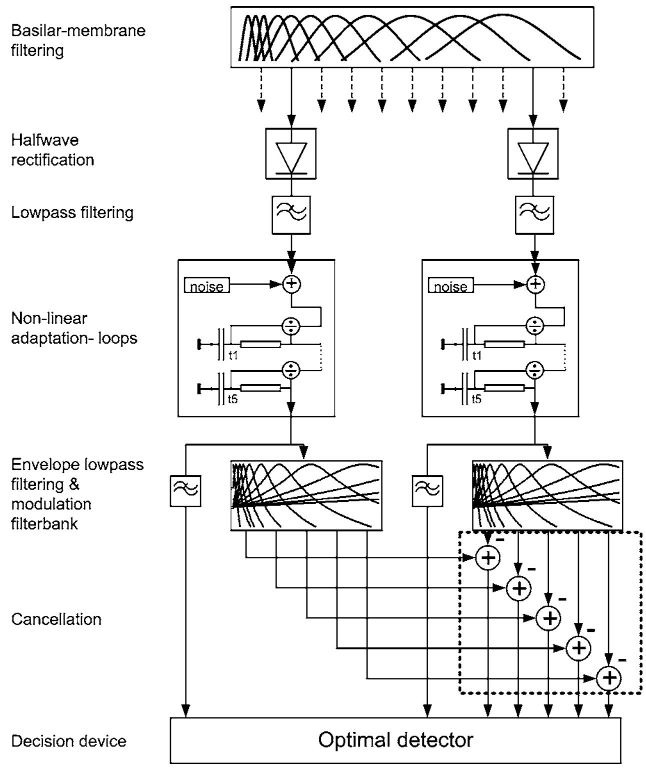

FIG. 1. Block diagram of the across-channel modulation filterbank model. The signals are filtered by the gammatone filterbank, half-wave rectified and low-pass filtered at 1 kHz, and subjected to adaptation. The adapted signal is then filtered by a modulation bandpass filterbank and a separate low-pass filter (at 2.5 Hz) at the output of each auditory filter. At the output of the individual modulation bandpass filters, the activity at the flanking bands is averaged across the flankers (E-process) and subtracted from the corresponding activity at the on-frequency band (C-process), illustrated here with only one flanking band and highlighted in the dashed box. The output activity is added to internal noise and finally subjected to an optimal detector as decision device.

larly to a power spectrum model (e.g., Patterson and Moore, 1996) and would account for certain aspects of spectral masking data (Derleth and Dau, 2000). In the second path, the bank of modulation bandpass filters captures the dynamic properties of the stimulus. It is expected that, in the model, a hypothetical process underlying across-channel CMR would use the output of the bandpass modulation filters. So far, however, the model in its original form does not contain any explicit across-channel processing and therefore fails to produce "true" across-channel CMR.

The present study introduces an explicit across-channel mechanism into the model. Figure 1 illustrates the model used in the present study. The modification of the model is comparable to the EC mechanism of Durlach's model (Durlach, 1960, 1963) for describing binaural masking level differences (BMLDs). However, while the EC mechanism in the original (binaural) model is applied essentially to the stimulus wave forms, and jitter is provided in the level and time domains in order to limit the resolution in the model, the (monaural) EC process in the current model is applied at a much later stage of auditory processing, and no additional limitations are introduced. In contrast to the original binaural EC model, it is assumed here that the limitations for performance are already included in the processing stages prior to the EC process.

The essential aspects of this approach are first illustrated for only two peripheral channels, i.e., using a channel centered at the on-frequency band including the signal, and a channel centered at one remote flanking band.

The across-channel processing within the model is assumed to occur at the output of all (bandpass) modulation channels tuned to frequencies at and above 5 Hz, which is the center frequency of the lowest modulation filter. The individual modulation filter outputs at the flanking band are subtracted from the corresponding outputs at the on-frequency channel. This process is denoted as cancellation in Fig. 1. The outputs of the low-pass filters in the different peripheral channels remain unaffected. The low-pass filtered outputs as well as the difference representations after modulation bandpass filtering are subjected to the decision stage, the optimal detector, which assumes independent observations for the different inputs. The specific case with only one flanking band does not require an equalization stage.

Typically, more than one flanking band will be presented. The generalized mechanism for the multi-channel case is indicated in Fig. 2. Here, the weighted sum of the activity of the flanking bands is computed and subtracted from the on-frequency channel. Calculating the weighted sum can be considered as equalization process, since it equalizes the summed activity in the different flanking bands with regard to the on-frequency band. The subtraction refers to a cancellation process as in the case with only one flanking band (Fig. 1).

In Fig. 2, a situation with N flanking bands and one on-frequency band is assumed. Here, the EC mechanism acts on the N peripheral channels, denoted as PC1, PC2, …, PCN. PCX indicates the channel centered at the on-frequency band. For simplicity, only the output of the jth modulation filter $s_{jn}(t)$ in the different peripheral channels ($n=1\ldots N$) is indicated in the figure. The outputs of all other modulation filters are processed in the same way. The output $s_j(t)$ of the EC mechanism for N channels at the jth modulation filter can be expressed as

$$s_j(t) = s_{jx}(t) - c_j(t) = s_{jx}(t) - \frac{\sum_{i=1,\ i\neq x}^{N} w_i a_i s_{ji}}{\sum_{i=1,\ i\neq x}^{N} w_i a_i}, \qquad (1)$$

where the index x denotes the peripheral channel (PCX) tuned to the on-frequency band and $c_j(t)$ represents the cancellation term. The contributions $s_{j1}, s_{j2}, \ldots, s_{jN}$ are weighted by the factors $a_1, a_2, \ldots, a_N$. The weights a_i equal the root mean square (rms) of the low-pass filter output in the channels PCi ($i=1,\ldots,N$). Since the rms value reflects the average energy of a signal, a_i equals the average energy in the ith peripheral channel. Thus, the weighting with a_i means that the channels that are excited by more input stimulus energy are emphasized relative to the filters which are excited by less. Specifically, filters without excitation by the stimulus do not contribute at all to the cancellation term $c_j(t)$. The cancellation term includes a normalization by the factor $\Sigma_{i\neq x,\ i=1}^{N} w_i a_i$ that is proportional to the overall energy of the stimuli in all peripheral channels except the on-frequency channel. In order to make sure that the EC stage operates *across* channels and does not subtract much

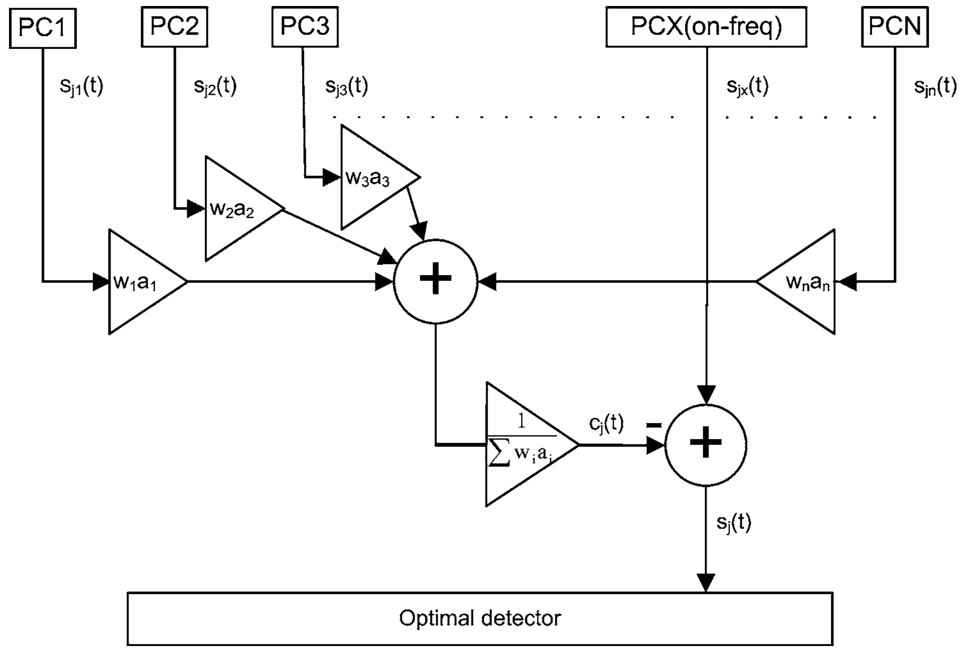

FIG. 2. Simplified block diagram of the across-channel EC process in the perceptual model for N peripheral channels PC1, …, PCN. Only one modulation filter at each peripheral channel is shown.

signal information from the signal channel, the off-frequency weight w_i was introduced. In the current implementation, w_i was set to zero if the overlap between the magnitude transfer function of the auditory channels at *PCi* and *PCX* is above a certain limit, and was set to one otherwise. The overlap of the filter transfer functions was calculated during the design phase of the model as the correlation value of broadband noise at the output of the two respective filters. The limit was chosen to be a correlation value of 5%. In this way, auditory filters tuned at ($i=x$) and very close ($w_i=0$) to the signal frequency were not considered in the EC process. The weights w_i ensure that, for example, in the case of a broadband noise as input, the stimuli in the channels contributing to the cancellation term are statistically independent from the excitatory on-frequency channel. Thus the EC mechanism in the model can be regarded as a true across-channel process.

In the most general version of the model, the EC process would be considered in all peripheral channels covering the whole audible frequency range, with each of the channels being regarded as a potential signal channel and with all respective surrounding channels being included in the cancellation term. In the simulations of the present study, however, the model was "told" in advance which was the signal frequency and thus which was the on-frequency channel. All remaining channels in the range from 500 to 6000 Hz were considered as the cancellation channels. This simplification is based on the assumption that the best signal-to-noise ratio is expected to be in the channel tuned to the signal and that detection is mainly based on this single channel (including the information from the other channels contained in the cancellation term of the EC process). An additional simplification was made in conditions when the stimulus was sparsely represented along the peripheral channels as, e.g., in the case of widely spaced narrowband flankers in first experiment. In this case, only channels tuned to the frequencies of the flanker bands were considered. The off-frequency weights w_i were then equal to one. If all flanker bands have equal energy (as in Experiment 1), all a_i have the same value a. The cancellation term $c_j(t)$ in Eq. (1) can then be simplified to

$$c_j(t) = \frac{\sum_{i=1,\ i\neq x}^{N} a s_{ji}}{\sum_{i=1,\ i\neq x}^{N} a} = \frac{\sum_{i=1,\ i\neq x}^{N} s_{ji}}{N-1} \qquad (2)$$

and becomes the average over the number of flanking bands.

III. METHOD

A. Subjects

Four normal-hearing subjects participated in each experiment. Their ages ranged from 23 to 41 years. All subjects had experience in other psychoacoustic experiments. The authors T.P. and T.D. participated in the experiment. The other two subjects were paid for their participation on an hourly basis.

B. Apparatus and stimuli

The subjects were seated in a double-walled, sound attenuating booth and listened via Sennheiser HD580 headphones. Signal generation and presentation during the experiments were computer controlled using the AFC software package for MATLAB, developed at Universität Oldenburg and DTU. All stimuli were generated digitally on an IBM compatible PC and were then converted to analog signals by a high-quality 32 bit soundcard (RME DIGI-96PAD) at a sampling rate of 32 kHz. Three CMR experiments were performed where the subject's task was to detect a tone in the presence of one or more noise masker bands. The specific stimuli will be described in the respective experiments (Experiments 1–3).

C. Procedure

A three-interval, three-alternative forced-choice paradigm was used to measure detection thresholds. A two-down, one-up procedure was used to estimate the 70.7% correct point of the psychometric function (Levitt, 1971) Subjects had to identify the one randomly chosen interval containing the signal. Subjects received visual feedback if the response was correct. The three observation intervals were separated by 500 ms of silence. The initial step size for the signal level was 4 dB and after every second reversal of the level adjustment the step size was halved until the step size of 1 dB was reached. The mean of the signal level at the last six reversals was calculated and regarded as the masked threshold value. For each stimulus configuration and subject, four masked threshold values were measured. The mean of these values was calculated and taken as the final threshold. For the model predictions, the identical procedure and the same alternative-forced-choice (AFC) framework as in the experiments were used.

IV. EXPERIMENT 1: CMR WITH FOUR FLANKING BANDS

A. Rationale

The first experiment was designed to investigate true across-channel CMR, where within-channel processing does not contribute. Four flanking bands with a spectral separation of one octave were used such that within-channel contributions to CMR can be assumed to be negligible at the (medium) sound pressure levels used in this experiment.

B. Stimuli

The signal was a 1000–Hz pure tone. The masker consisted of five bands of noise which were centered at 250, 500, 1000, 2000 and 4000 Hz, thus covering a frequency range of four octaves. Signal and masker had the same duration of 187.5 ms; 20–ms raised-cosine ramps were applied to the stimuli. Signal threshold was measured as a function of the bandwidth of the masker, which was 25, 50, 100 or 200 Hz. The masker bands were generated in the time domain, transformed to the frequency domain by Fourier transform where they were restricted to the desired bandwidth, and finally transformed back to the time domain by inverse Fourier transform. In the reference condition, the envelopes of the five bands were uncorrelated with each other. In the comodulated condition, the on-frequency noise masker was shifted to the center frequencies of the flanking bands in the Fourier domain, such that the envelopes of the different bands were fully correlated with each other. The presentation level of each of the maskers was 60 dB sound pressure level (SPL).

C. Results

Figure 3 shows the results of the experiment. Masked thresholds are plotted as a function of the masker bandwidth. The open symbols represent the experimental data, averaged across subjects. The circles and squares show the results for the uncorrelated and comodulated conditions, respectively. The right panel of Fig. 3 shows the amount of CMR, i.e., the difference between the uncorrelated and comodulated thresholds. There is a significant CMR effect of 4–5 dB for the

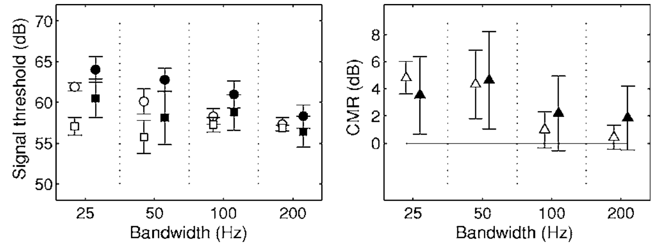

FIG. 3. Left panel: Detection thresholds for the 1–kHz tone in the presence of five noise bands as a function of the bandwidth of the noises. Open symbols indicate average experimental data and filled symbols show simulation results. Circles and squares represent results for the uncorrelated and comodulated conditions, respectively. Right panel: CMR effect for the conditions of the left panel.

small noise bandwidths of 25 and 50 Hz [one-way analysis of variance (ANOVA): $F(1,22)=38.59$, $p<0.001$ and $F(1,22)=32.18$, $p<0.001$], while no significant CMR was found for the larger bandwidths of 100 and 200 Hz [one-way ANOVA: $F(1,22)=1.67$, $p=0.21$ and $F(1,22)=0.02$, $p=0.89$] where statistical significance here and in the following is defined as having $p<0.01$. Thus, even though four flanking bands were used, the obtained CMR is relatively small compared to the results typically found with narrow spacing between the signal and flanking bands (see Experiment 2) or in the band-widening CMR paradigm (see Experiment 3). The results are consistent with results from previous studies (e.g., Moore and Emmerich, 1990), showing that CMR is restricted to narrowband noises with bandwidth smaller than 50 Hz. This indicates that across-channel CMR is a phenomenon that occurs only when the masker is dominated by relatively slow envelope fluctuations. The modulation spectrum of bandpass noise is directly related to the bandwidth of the noise (e.g., Lawson and Uhlenbeck, 1950; Dau *et al.*, 1997a). The rate of modulations will range up to the bandwidth of the noise, Δf.

The filled symbols in Fig. 3 show the simulations obtained with the processing model described in Sec. II. The simulations represent average thresholds of ten repetitions for each experimental condition. The model predicts slightly elevated overall thresholds (2–3 dB) and larger standard deviations in comparison to the empirical data. For the bandwidths 25 and 50 Hz, the model predicts a significant mean CMR effect of about 4 dB [one-way ANOVA: $F(1,18)=15.38$, $p<0.001$ and $F(1,18)=16.91$, $p<0.001$, respectively]. It does not produce a significant amount of CMR for the 100 and 200 Hz bandwidths [one-way ANOVA: $F(1,18)=6.48$, $p=0.02$ and $F(1,18)=6.29$, $p=0.02$].

D. Model analysis

The following describes how the EC mechanism affects the signal processing of the stimuli in the model. Since the EC process typically leads to a lower threshold in the comodulated condition compared to the uncorrelated condition, this should be reflected in the model's internal representations of the stimuli. As an example, the upper left panel of Fig. 4 shows the internal representation of a single 25-Hz-wide (comodulated) noise masker centered at 1 kHz. The outputs of the modulation filters are shown separately in

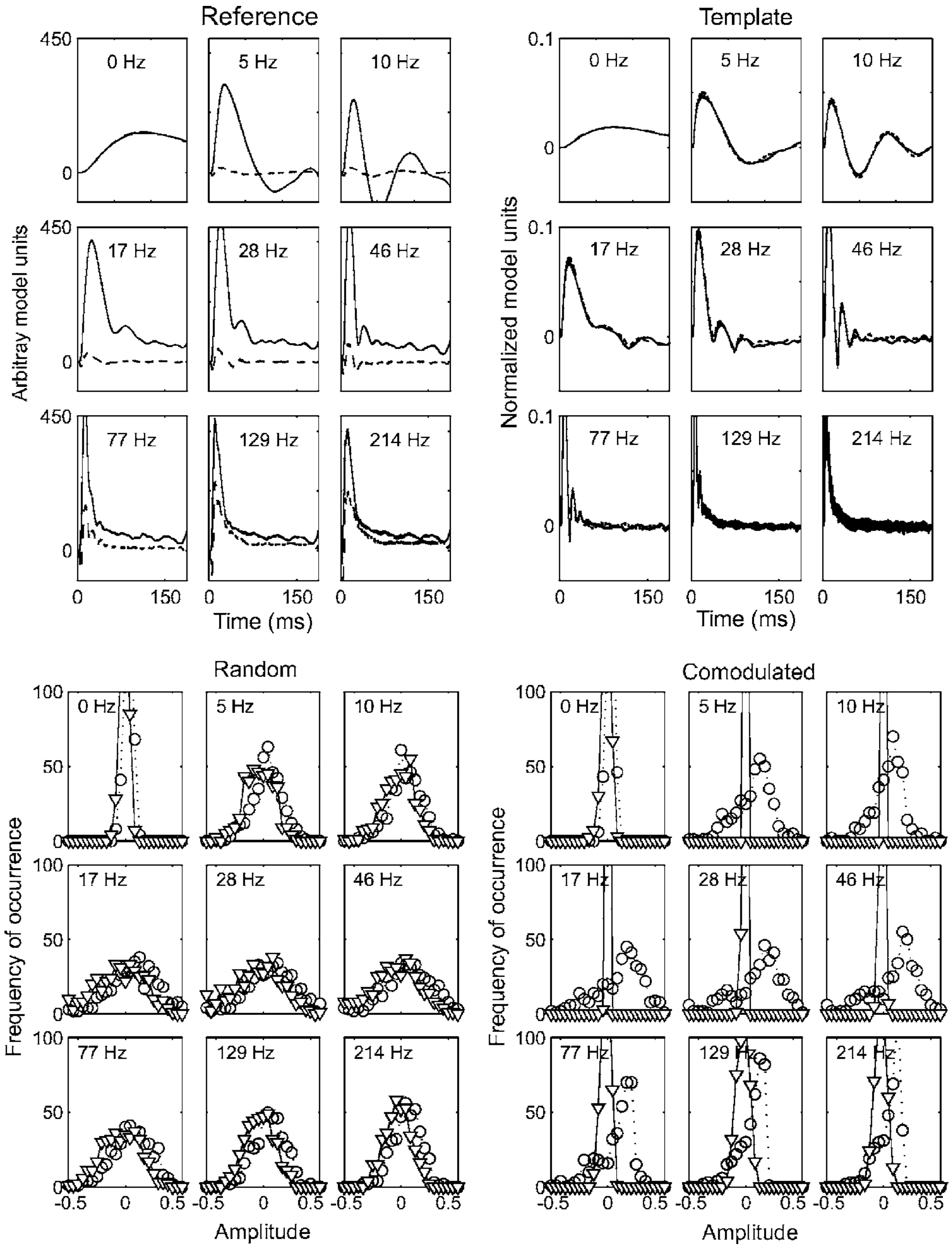

FIG. 4. Simulated internal representations at the output of the modulation filters (indicated by the center frequencies in the subpanels) in the on-frequency (peripheral) channel. Solid curves show outputs without EC process, dashed curves show results after the EC process. Left upper panel: Internal representation of (modulated) noise alone (i.e., no signal was added). Right upper panel: Internal representation of the template, i.e., the normalized difference between noise plus supra-threshold signal representation and noise alone representation. The lower panels show histograms of the cross-correlation coefficients between the noise-alone representation and template (triangles, solid line), and between the noise-plus-actual-signal representation and template (circles, dotted line), for the same individual modulation filters as considered in the top panels. This is shown for the random condition (left) and the comodulated condition (right), with EC mechanism applied.

the subpanels, including the modulation low-pass filter (indicated as 0 Hz), and the bandpass filters tuned to 5, 10, 17, 28, 46, 77, 129, and 214 Hz. The solid curves show the output obtained without EC process, i.e., when using the original model's (Dau *et al.*, 1997b) preprocessing. The dashed curves show the output when the EC process was included, i.e., after subtracting the average activity of the four flanking bands from the on-frequency band. As expected, the output representation (for the comodulated noise bands) after the EC process is reduced in amplitude compared to the result without the EC process. Note that modulation channels tuned to frequencies higher than the bandwidth of the noise (25 Hz) are activated as well, mainly reflecting the response to the onset of the adapted envelope of the stimulus.

As described in Sec. II and in previous publications (Dau *et al.*, 1996, 1997a), in the simulations, the internal representation of the noise is subtracted from the internal representation (either noise alone or signal plus noise) of each of the three intervals and then cross correlated with the template. The template represents the normalized difference between the internal representation of the noise plus *supra*-threshold signal and the noise-alone representation. The upper right panel of Fig. 4 shows the model's template using the same 25-Hz-wide noise (as used for the illustration of the reference) to which a supra-threshold 1–kHz tone was added. As for the reference representation, the individual modulation filter outputs are indicated in the subpanels. In the case of the template, there is essentially no difference between the situation with and without EC process since the internal representation of the template is dominated by the presence of the signal.

In order to evaluate the function of the EC mechanism, the two lower panels of Fig. 4 show a statistical analysis of the cross correlation between noise-alone representation and template (triangles), and between noise-plus-actual-signal representation and template (circles) including the EC mechanism in the processing. The histograms of the cross-correlation coefficient are shown for the output of the same individual modulation filters as considered in the top panels. The "actual" signal level was chosen to be the simulated signal level at threshold (from Fig. 3, random condition). For the template, the same supra-threshold level (85 dB) was used as in the simulations. The lower left panel shows the results for the random noise condition at the output of the EC process. Since the signal level was chosen to be at detection threshold, the distributions are just separable (in terms of signal detection theory). The right panel shows the corresponding results for the comodulated condition. Here, the EC mechanism causes a strong sharpening of the distribution of correlations in the reference interval while the distributions in the signal interval remain essentially unaffected. This corresponds to an increased sensitivity and a decreased detection threshold in the simulations in the comodulated condition relative to the random condition, and represents the "noise reduction" caused by the EC mechanism. Without the EC mechanism, the histograms are similar in the random and comodulated condition.

The comparison of the histograms at the output of the different modulation filters suggests that all modulation filters contribute to signal detection (also those tuned to modulation frequencies higher than the noise bandwidth of 25 Hz). In other words, the decision in the model does not seem to be based on the activity at the output of only one or a few particular modulation filters. This is different from the situation in conditions of within-channel CMR, at least in the framework of the current model, where modulation cues like beatings between on-frequency and flanker bands components become effective and activate specific modulation filters in the signal interval (see the corresponding analysis in Experiment 2). In the EC model, a supra-threshold signal does not produce any specific modulation pattern that could be used as cue. The EC mechanism therefore does not lead to an enhancement of specific cues which would be reflected by different templates for the same condition with or without EC mechanism. The EC mechanism rather suppresses the noise fluctuations in the modulation filters, thereby enhancing signal detection.

Since the outputs of all bandpass modulation filters contribute to the function of the EC mechanism in the model, the question remains whether a modulation filterbank is necessary for the occurrence of CMR. To address this question, additional simulations were carried out with alternative modulation filtering stages: (i) A process referred to as "DC/AC" which separates the dc component of the Hilbert envelope spectrum from the remaining (ac) spectrum, (ii) a combination of a second-order Butterworth low-pass and a high-pass filter with cutoff frequencies of 2.5 Hz, referred to as "LH," (iii) a combination of the same low-pass filter at 2.5 Hz combined with a single bandpass filter centered at 5 Hz with a bandwidth of 5 Hz, referred to as process "LB5," and (iv) the same as (iii) but with a bandpass filter tuned to 50 Hz and a bandwidth of 25 Hz ($Q=2$; referred to as "LB50"). The EC process was applied to the ac-coupled output in dc/ac, the output of the high-pass filter in LH, and the output of the single bandpass filters in LB5 and LB50,

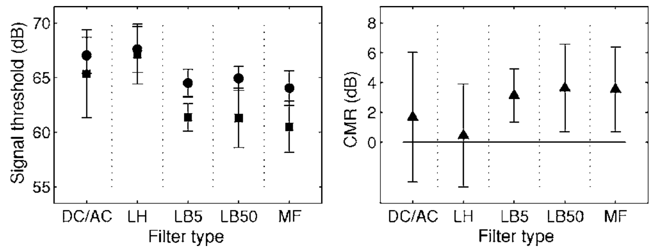

FIG. 5. Left: Signal thresholds obtained with the filter types dc/ac, LH, LB5, LB50, as defined in the main text and the complete modulation filterbank (MF). Circles and squares show results for random and comodulated noise, respectively. Right: Amount of CMR for the different filter types.

respectively. Figure 5 (left panel) shows the corresponding predictions obtained with the different processing schemes for the random and the comodulated noise conditions using the same symbols as in Fig. 3. The right panel shows the amount of CMR for the different schemes. The result obtained with the complete modulation filterbank, referred to as "MF," was replotted from Fig. 3 for direct comparison.

The DC/AC and LH processes do not produce any CMR [one-way ANOVA: $F(1,18)=1.51$, $p=0.24$ for DC/AC, $F(1,18)=0.16$, $p=0.68$ for LH]. In contrast, the processing of LB5 and LB50 produces a significant CMR effect of about 4 dB [one-way ANOVA: $F(1,18)=30.96$, $p<0.001$ and $F(1,18)=15.4$, $p<0.001$] which corresponds to the prediction obtained with the complete modulation filterbank MF [one-way ANOVA: $F(1,18)=38.59$, $p<0.001$]. Thus, within the model, across-channel CMR can only be produced if the stimulus after peripheral filtering, envelope extraction and adaptation is actually processed by frequency-selective (modulation) filters, whereby each individual filter would already be sufficient to produce significant CMR. The effect, however, disappears if only one broad (5–150 Hz) modulation bandpass filter is considered (not shown explicitly). The reason for the behavior in the model is that the input to the modulation filtering process, the adapted envelope, shows an onset response. This onset produces an excitation also at higher modulation frequencies. The EC process is only effective if the output of the modulation filtering process leads to a reasonable correlation between the flanking band and the signal band representations. This is only the case after (modulation) bandpass filtering, and cannot be obtained for the "broadband" schemes DC/AC and LH considered above. It is not clear, of course, to what extent the mechanisms in the real system are related to the ones proposed here on the basis of the model. The intention of the above analysis was to elucidate the functioning of the EC process of the proposed model.

In summary, the data from Experiment 1 confirm results from previous studies that across-channel processing in CMR is robust but small (even when several flanking bands are involved). Across-channel CMR is only observable at small bandwidths (below about 50 Hz), i.e., when the envelope fluctuations inherent in the stimuli are relatively slow. The simulations indicate that across-channel CMR can be accounted for quantitatively if an EC-like mechanism is introduced at the output of a modulation frequency-selective process.

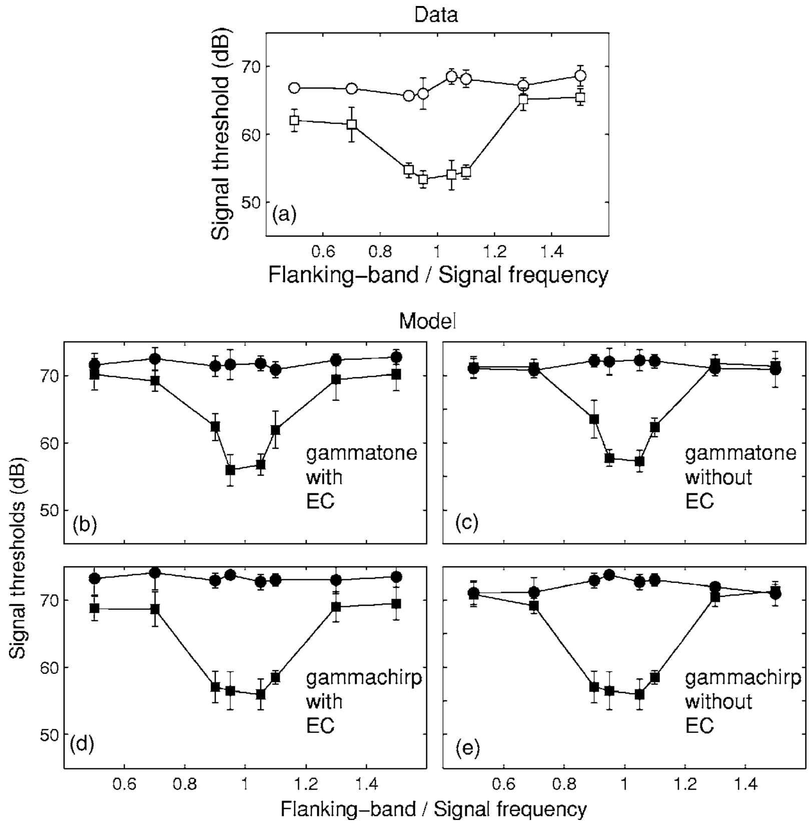

FIG. 6. (a) Measured data averaged across subjects. Signal threshold for a 2–kHz tone in 25–Hz wide noise as a function of the spectral separation between on-frequency band and flanking band. Circles and squares show results for random and comodulated noise, respectively. (b) Predictions with the EC model shown in Fig. 1, using gammatone filters as the peripheral filtering stage. (c) Predictions with the same model, but with EC process switched off. (d) Predictions as in (b) but with gammachirp filters. (e) Predictions with gammachirp filters, but with EC process switched off.

V. EXPERIMENT 2: CMR WITH ONE FLANKING BAND VARYING IN FREQUENCY

A. Rationale

This experiment investigates the transition between conditions where exclusively across-channel mechanisms determine CMR and those where primarily within-channel mechanisms generate CMR. Only one flanking band was used here, as in the study by Schooneveldt and Moore (1987). The amount of CMR was measured and simulated as a function of the spectral separation between the flanking and the on-frequency band. While for large separations of one octave or greater, CMR cannot be expected to exceed 2–4 dB, masking releases of about 14 dB and higher were observed in previous studies for separations of less than 1/10 octave where within-channel processing provides the most effective detection cues (Schooneveldt and Moore, 1987). A successful model of CMR needs to account for both within- and across-channel components.

B. Stimuli

The stimuli were similar to some of those used in Schooneveldt and Moore (1987). The signal was a 2000 Hz tone. The on-frequency masker was a 25-Hz-wide band of noise centered at the signal frequency. The flanking band had the same bandwidth as the on-frequency band and was centered at 1000, 1400, 1800, 1900, 2100, 2200, 2600 or 3000 Hz, corresponding to frequency ratios between flanking band and on-frequency band of 0.5, 0.7, 0.9, 0.95, 1.05, 1.1, 1.3, and 1.5. In contrast to the study by Schooneveldt and Moore (1987), the flanking band was not presented directly at the signal frequency or very close to it. The two noise bands were either uncorrelated or comodulated. As in Schooneveldt and Moore (1987) each band was produced by multiplying a sinusoid at the center frequency with a low-pass noise with a cutoff frequency of 12.5 Hz. In the comodulated condition, the noise bands were produced by multiplying the different sinusoids with an identical low-pass noise whereby a new noise was generated for each interval. Each band had an overall level of 67 dB SPL.

C. Results and model analysis

Panel (a) of Fig. 6 shows average data for the uncorrelated (open circles) and the comodulated (open squares) conditions. The signal threshold is plotted as a function of the ratio between flanking-band and signal frequency. The difference in threshold between uncorrelated and comodulated conditions, i.e., the amount of CMR, reaches 12–14 dB when flanker and signal frequency are close to each other (with ratios between 0.9 and 1.1). For large separations between on-frequency and flanking band, the data show a slight asymmetry: CMR of 3–4 dB in the presence of the high-frequency flankers and 5–6 dB for flanking bands presented at low frequencies. The data agree well with the results of Schooneveldt and Moore (1987).

Panel (b) of Fig. 6 shows the simulations obtained with the present model. As described in Sec. II, the EC mechanism was applied in all filters that overlap less than 5% with the on-frequency gammatone filter, i.e., in all channels except the two closest ones on both sides of the on-frequency channel. In this particular experiment, this means that the flanker bands were maximally contributing to the cancellation term of the EC process at frequency ratios of 0.5, 0.7, 1.3, and 1.5. The model accounts for the relatively flat threshold function obtained in the uncorrelated condition. For flanking-band frequencies close to the signal frequency (at the frequency ratios 0.95 and 1.05), the model predicts a large amount of CMR that corresponds to that found in the experimental data. This component depends on beating of the carrier frequencies of the on-frequency and flanking bands. In the model this can be accounted for by the processing within the (peripheral) channel tuned to the signal frequency. The model detects changes in the envelope statistic due to the addition of the signal to the on-frequency band (see Verhey *et al.*, 1999). This is effective for the comodulated condition while it does not provide additional detection cues in the uncorrelated condition. At very low and very high flanking band frequencies, the model predicts an average amount of CMR of about 3 dB which agrees well with the data at the high flanker frequencies but is slightly less than the measured effect at the low flanker frequencies. The simulated 3 dB effect is the result of the EC mechanism in the model as can be seen from direct comparison with the results obtained without EC circuit, shown in panel (c) of Fig. 6. As expected, without across-channel processing, no CMR is predicted at the large frequency separations between the on-frequency and the flanker band.

While certain aspects of the data can be described satisfactorily by the model, some other aspects cannot. First, the simulated threshold function for the comodulated condition increases too steeply with increasing spectral distance from the signal. Second, the predicted amount of CMR for the lowest flanker frequencies is smaller than in the data. The reason for these discrepancies might be related to the shape of the magnitude transfer function of the peripheral filters used in the simulations. The gammatone filters are symmetrical on a linear frequency scale. However, it has been demonstrated that below its center frequency, the skirt of the human auditory filter broadens substantially with increasing stimulus level, and above its center frequency the skirt sharpens slightly with increasing level (Lutfi and Patterson, 1984; Moore and Glasberg, 1987). In order to illustrate effects of frequency selectivity on CMR in the framework of the current model, additional simulations were carried out using gammachirp filters (Irino and Patterson, 1997). The gammachirp filter has an asymmetric magnitude transfer function, and the degree of asymmetry in this filter is associated with stimulus level. The gammachirp filter was shown to provide a very good fit to human notched-noise masking data. Its impulse response is well defined and includes only one parameter more than the gammatone filter (see Eq. 2, in Irino and Patterson, 1997). In the present study, the impulse responses of the gammachirp filters were calculated for a level of 67 dB SPL. Here, as a simplification, the simulations were run with selected gammachirp filters tuned to the on-frequency band and the flanking band, respectively. A complete gammachirp filterbank with well defined level-dependent overlap has not been developed yet. As in the previous simulations with gammatone filters, the EC process was applied when the overlap between the off-frequency channel and the signal channel was below 5% which was only the case for the four outer data points (frequency ratios 0.5, 0.7, 1.3, and 1.5). All other model parameters were kept the same as in the simulations with gammatone filters. The results are shown in panel (d) of Fig. 6.

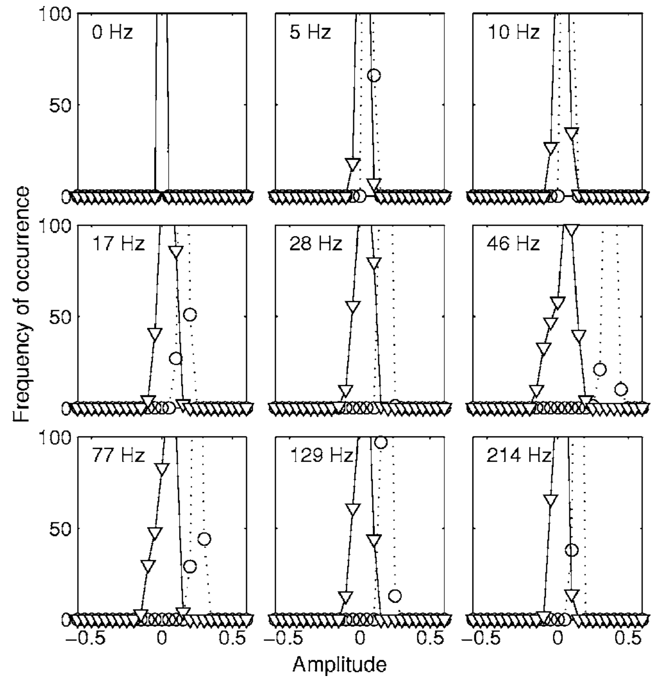

FIG. 7. Histograms of the cross-correlation coefficients at the output of nine modulation filters in an exemplary condition of Experiment 2 with 50 Hz separation between the on-frequency and flanking bands. Correlations for the reference (triangles, solid line) and reference plus signal (circles, dashed line) are shown for comodulated noise bands. The center frequency of the modulation filter is indicated within each panel. For the output of the modulation filter close to 50 Hz, the mean of the distributions is most different and the distributions are most separable in terms of signal detection.

The simulations with gammachirp filters account for many aspects of the experimental data. Due to the broader bandwidth of the gammachirp filter compared to the gammatone filter, within-channel cues become effective for a larger range of flanking-band frequencies. The plateau of low thresholds corresponds to that found in the data. At low and at high flanking-band frequencies, CMR amounts to 3–4 dB due to the EC processing in the model. However, the introduction of the gammachirp filter does not account for the slight asymmetry observed in the measured data, even though the transfer functions of the individual filters have an asymmetric shape. The simulated pattern for the comodulated condition actually produces the same thresholds at both ends. Still, the overall correspondence with the data is high. For direct comparison, panel (e) of the same figure shows the corresponding predictions without EC process. All data points except for the four outer ones are replotted from panel (d), since no EC process was applied for the inner data points in panel (d). As for the simulations with gammatone filters without the EC process, no CMR was obtained at the largest spectral separations between flanking and on-frequency band.

In order to illustrate the importance of within-channel cues available in the conditions where on-frequency band and flanking band are close to each other, Fig. 7 shows a statistical analysis similar to that presented in Experiment 1.

Histograms of the cross correlation between noise-alone representation and template (triangles) and noise-plus-actual-signal representation and template (circles) are shown for the outputs of the individual modulation filters. An exemplary frequency separation between on-frequency band and flanking band of 50 Hz was used for illustration. It can be seen in Fig. 7 that signal detection is mainly based on information at the output of the modulation filter tuned to about 46 Hz. Here, the mean of the signal distribution is clearly larger than that of the noise distribution. Thus, the addition of the signal to the masker causes changes in the internal representation of the stimuli such that it can effectively be evaluated in one (or only a few) modulation filters in this given task. This detection cue is qualitatively different from that discussed in connection with the across-channel process where signal detection was mainly based on the sharpening of the noise distribution at the output of the EC process in all modulation filters.

The results of Experiment 2 thus support the hypothesis that CMR has (at least) two components. One is restricted to flanking band frequencies around the signal frequency. This component reflects the use of within-channel cues (beating), rather than across-channel cues. The other component does not depend strongly on flanking-band frequency, but rather on across-channel cues. This across-channel component of CMR amounts to about 3 dB. While this has been proposed in earlier studies (e.g., Schooneveldt and More, 1987), the present study tried to provide quantitative modeling to test explicitly the (relative) contributions of within- and across-channel processing.

VI. EXPERIMENT 3: CMR AS A FUNCTION OF THE MASKER BANDWIDTH

A. Rationale

The third experiment considered the “classical” band-widening experiment where the masker was centered at the signal frequency and signal threshold was obtained as a function of the bandwidth of the masker. In contrast to the two previous experiments, the band-widening experiment does not allow for a separation between within- and across-channel processes; within-channel contributions will always contribute to CMR, even for large masker bandwidths when many auditory filters are excited by the noise. Verhey *et al.* (1999) showed that a single-channel analysis, which uses only the information in one peripheral channel tuned to the signal frequency, quantitatively accounts for the main CMR effect in the band-widening experiment. This suggested that across-channel processes are not involved or not effective in this class of CMR experiments, even though several auditory filters are excited by the noise. This was directly investigated here with the extended model that includes an explicit across-channel process while it keeps the ability to process within-channel cues, as shown in Experiment 2.

B. Stimuli

The signal was a 300-ms-long, 2000–Hz pure tone. The masker was a band-limited noise centered at the signal frequency. The masker bandwidth was 50, 100, 200, 400, 1000

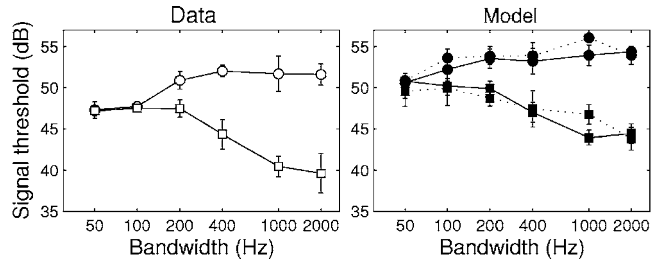

FIG. 8. Left panel: Average signal thresholds for four subjects are plotted as a function of the masker bandwidth in random noise (circles) and comodulated noise (squares). Right panel: Predicted signal threshold of the model when the EC mechanism is applied (dashed line) and when it is not applied (solid line). The modulator bandwidth was 50 Hz and the signal frequency was 2000 Hz.

or 2000 Hz. The duration of the masker was 600 ms with 10–ms raised-cosine onset and offset ramps. The signal was temporally centered in the masker. Two types of maskers were used, as in the original experiments in Hall *et al.* (1984a). One was a random noise with irregular and independent envelope fluctuations in different frequency regions. The comodulated noise was a random broadband noise which was modulated in amplitude at an irregular, low rate, and then restricted to the desired bandwidth. A low-pass noise with a cutoff at 50 Hz was used as a modulator. Other studies have shown that for modulator bandwidths larger than 50 Hz, CMR decreases with increasing modulator bandwidth whereas it remains roughly constant for modulator bandwidth, below this value (Schooneveldt and Moore, 1987; Carlyon and Stubbs, 1989). The modulation resulted in fluctuations in the amplitude of the noise which were the same in different frequency regions. The spectrum level of the bandpass noise was 30 dB, corresponding to overall levels of 47–63 dB SPL for the 50–2000 Hz bandwidth range.

C. Results

Figure 8 shows the results of the band-widening experiment. The left panel shows the experimental data, averaged across subjects. The signal threshold is plotted as a function of the masker bandwidth, for random noise (open circles) and comodulated noise (open squares). Consistent with the results from the earlier studies, for the random noise, the masked threshold first increases as the masker bandwidth is increased. Beyond a certain bandwidth (200 Hz in this case), the threshold no longer increases, but remains roughly constant. The increase of the threshold is caused by the fact that, up to the critical bandwidth, more noise passes through the auditory filter centered at the signal frequency, while beyond the critical bandwidth, the added noise falls outside the passband of the auditory filter. In contrast, for the comodulated noise, the threshold first stays constant and then decreases as the bandwidth is increased beyond about 200 Hz. The amount of CMR, defined as the difference in threshold between the random and comodulated conditions, is 12 dB for the largest bandwidth (2000 Hz).

The right panel of Fig. 8 shows the corresponding model predictions. For direct comparison, simulations are shown with EC process (dashed line) and without EC process (solid line). The two model versions essentially produce the same

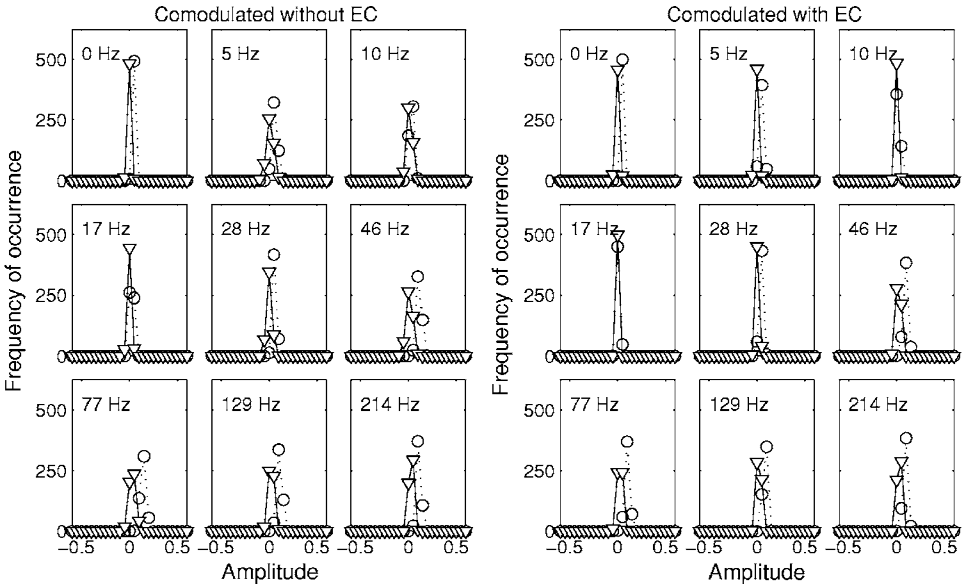

FIG. 9. Histograms of the cross-correlation coefficients at the output of nine modulation filters for the comodulated conditions of Experiment 3 with a noise bandwidth of 2000 Hz. Left panel: Reference alone (triangles, solid line) and reference plus signal (circles, dashed line) for comodulated noises without EC process. Right panel: Reference alone (triangles, solid line) and reference plus signal (circles, dashed line) for comodulated noises with EC process included.

results. Thus, the across-channel processing does not generate any change in the overall amount of CMR in the framework of this model, not even at the largest masker bandwidths where several auditory channels are excited. Figure 9 shows the statistical analysis of the decision variable in the simulations, as in the first two experiments.

The comodulated condition with the broadest noise bandwidth (2000 Hz) was considered for illustration with and without EC mechanism. At this bandwidth, the observed amount of CMR is maximal (12 dB). The analysis was carried out at a signal level of 55 dB which is about 10 dB above the simulated threshold in the comodulated condition. The left panel shows the distribution of the cross-correlation between noise-alone representation and template (triangles) and for the signal-plus-noise representation and template (circles) at the output of the single (peripheral) channel tuned to the signal frequency (single-channel analysis). It can be seen that there is a separation between the two distributions at several modulation filter outputs. Since the bandwidth of the noise (also after peripheral filtering) is larger in this experimental condition than in the previous experiments, the variability of the envelope amplitude fluctuations is smaller, leading to the relatively sharp distributions. The right panel shows the analysis including the across-channel process in the model, i.e., a multi-channel simulation was carried out in this case where the cancellation term in the EC process was derived from the off-frequency channels. The envelope correlation across the different peripheral channels is *not sufficient* to effectively increase the signal-to-noise ratio at the output of the EC process in the model. The EC process therefore does not contribute to signal detection in this type of experiment in the framework of the model.

These results therefore support the hypothesis that CMR obtained in the band-widening paradigm is strongly dominated by within-channel processing and is not a result of across-channel processing.

VII. OVERALL DISCUSSION

A. Within- versus across-channel processing

The modeling results of this study support the hypothesis that (at least) two mechanisms are contributing to what has been defined as CMR. The present model allows a distinction to be made between these two contributions. The simulations strongly support that one of the processes is based on within-channel mechanisms. Signal detection is based on the changes of the internal representation of the stimuli at the output of individual auditory filters—without the need for explicit across-frequency processing. The addition of the signal to the comodulated masker typically changes the (envelope) statistics of the stimuli significantly, while the changes are much smaller or absent in the case of random noise maskers (Schooneveldt and Moore, 1987; Verhey *et al.*, 1999). CMR resulting from within-channel contributions can be up to about 15 dB depending on the specific condition, and modulations (for example, resulting from beatings between signal and masker components) up to several hundred Hz can serve as a cue for signal detection. A pre-requisite for accounting for the full range of within-channel contributions to CMR is therefore a high sensitivity of the model to amplitude modulations (see also Verhey *et al.*, 1999), as is the case for the modulation filterbank used in the present framework. Specifically, the modeling results suggest that a few individual modulation filters (at the output of the single peripheral channel at the signal frequency) can process the changes in the internal representation of the stimuli effectively.

The other form of CMR is based on true across-channel processing. This effect is also robust but relatively small (2–4 dB) and becomes only effective when narrowband noises with bandwidths below about 50 Hz are presented, i.e., when the envelope fluctuations of the noises vary relatively slowly. The EC model described in the present study makes specific assumptions about how envelope information at the output of different auditory channels might be processed. The EC process was assumed to take place at the output of each modulation bandpass filter. The effect of the EC process is that the variance of the external noise (originating from the masker) at the level of the internal representations after the EC process is reduced in the case of the comodulated noise condition. This leads to improved signal detection compared to the random noise condition. In the framework of the model, the detection cue is thus qualitatively very different from the situation where within-channel processing determines CMR.

It is clear that effects of nonlinear peripheral processing, such as the level-dependent auditory filter bandwidth, have an influence on the relative contributions of within- and across-channel processing to CMR. In fact, some of the effects that were considered as across-channel contributions in the past might become within-channel contribution with proper modeling of nonlinear filters. For example, at very high stimulus levels where the auditory filter bandwidth is markedly increased (compared to the gammatone filters used in the present study), it can be expected that even in conditions with very broad spacing between the on-frequency band and the flanking band(s), CMR might be dominated by within-channel contributions. Ernst and Verhey (2005) have shown that CMR over ranges of three octaves can be modeled as a suppression effect in a nonlinear single-channel model, using the dual resonance nonlinear filter (DRNL) model (Meddis *et al.*, 2001). In some of the conditions in their study, however, the level of the off-frequency flanker was much higher (up to 60 dB) than that of the on-frequency band. Although their results are not directly comparable to the experimental conditions used in the present study, it can be assumed that with proper modeling of the nonlinear auditory filters even more signal configurations that have been considered as across channel in the past might reveal a within-channel contribution. Our current definition of when across-channel processing is applied to a particular filter is based on the amount of overlap of its transfer function with that of the signal channel. This definition might be general enough to also apply to filters of different or varying shapes and to nonlinear filters; the approach was successful when analyzing the results of Experiment 2, where individual gammachirp filters were considered. This, however, needs to be further investigated using a complete filterbank of filters with different shape or a nonlinear filterbank such as, e.g, a bank of DRNL filters or a gammachirp filterbank.

The observation that two conceptually different mechanisms define CMR is compatible with the results from studies on effects of auditory grouping on CMR (Grose and Hall, 1993; Dau *et al.*, 2005). When widely spaced flanking bands were used (as in the first experiment of the present study), CMR effects could be eliminated completely by introducing a gating asynchrony between the on-frequency masker and the flanking bands, by introducing precursor flanking bands, and by introducing following flanking bands. Due to the large spacing (and the relatively low presentation levels), only across-channel processes contributed to CMR. In contrast, using narrowly spaced flanking bands with 1/6-octave spacing (similar to the conditions with close frequency spacings in Experiment 2), CMR was not affected by any of the stimulus manipulations. It was therefore suggested that (i) the within-channel mechanisms in CMR might be peripheral (brainstem level or below) in nature and therefore not susceptible to manipulation by auditory grouping constraints, and that (ii) the "slower" across-channel processing that is strongly dependent on auditory grouping constraints, might be of more central origin (Dau *et al.*, 2005). The model investigated in the present study is not able to identify or extract auditory objects based on comodulation. A more advanced version of the model might apply basic concepts of computational auditory scene analysis (Bregman *et al.*, 1990), where the EC-process would be switched on or off depending on the current spectro-temporal acoustical context.

B. Correlation with physiological CMR results

Even though a large number of studies have investigated CMR from a psychophysical perspective, little is known of its underlying physiological mechanisms (see, e.g., Verhey *et al.*, 2003, for a review). A few studies have addressed physiological mechanisms of across-frequency processing by estimating signal-detection thresholds from the recordings of single- and multi-unit recordings in CMR-like paradigms. Several stages of processing along the auditory pathway were considered. Some studies intended to investigate across-channel processing but actually studied mostly within-channel cues due to the specific choice of the stimuli (e.g., Mott *et al.*, 1990). Nelken *et al.* (1999) investigated the response of neurons in the primary auditory cortex to noise of varying bandwidth. They found a correlate for CMR in the band-widening paradigm in the disruption of the neurons' envelope following response. For most of the neurons in the population, the envelope locking was degraded by the addition of the pure tone signal. Using statistical criteria to estimate signal detection threshold, Nelken *et al.* (1999) demonstrated that the suppression of envelope locking lowers the detection thresholds for the single tones when comparing the responses of modulated versus unmodulated noise bands.

When considering true across-channel CMR, two possible correlates have been discussed recently. In the primary auditory cortex (of the cat), Rotman *et al.* (2001) in another study used a stimulus centered on the best frequency of the neuron and added two flanking bands equally spaced at either side of the best frequency. They showed that a single unit in the auditory cortex can demonstrate a response consistent with CMR in the flanking band paradigm. The correlate of CMR was again found as a disruption of the envelope following response. Thus, it appears that CMR is coded at a relatively late stage of auditory processing (in the primary auditory cortex) which appears conceptually compatible with the psychophysical findings on grouping constraints on CMR. Their finding of very similar correlates for CMR in the two stimulus paradigms seems to differ from the modeling analysis discussed in the present study, which suggests very different mechanisms for the two processes.

A second physiological correlate of across-channel CMR has been suggested to be wideband inhibition at brainstem level (e.g., Pressnitzer *et al.*, 2001; Meddis *et al.*, 2002). Here, it has been suggested, based on physiological experiments with the flanking-band paradigm with deterministic maskers, that cochlear nucleus onset units provide wideband inhibition at the level of the brainstem onto narrowband units in the ventral cochlear nucleus, and that this wideband inhibition could provide a possible physiological basis for a potential EC model of CMR (for details about hypothetical neural circuits underlying CMR in the cochlear nucleus, see Pressnitzer *et al.*, 2001; Verhey *et al.*, 2003). A problem with such a neural correlate at the level of the brainstem might be the perceptual findings in the context of auditory grouping which make it unlikely that across-channel CMR can be accounted for by processing in the auditory brainstem and below.

A very promising way to fully understand the physiological mechanisms underlying CMR might be to study the correlation between neural responses and performance in the same species. Such an investigation was undertaken by Langemann and Klump (2001) and Nieder and Klump (2001) using the starling. Nieder and Klump (2001) investigated across-channel CMR with the flanking band paradigm, but used 100-Hz-wide on-frequency and flanking bands amplitude modulated at 10 Hz. They showed that neural detection threshold was lowest when the probe tone was positioned in a dip of the masker envelope. They concluded that their multi-unit recordings in the auditory forebrain of the starling can be compared to the behavioral results in the same species. It would be interesting to specifically study the three basic paradigms of the present study in the same animal model both behaviorally and physiologically to learn more about the potential correlates of the different mechanisms underlying CMR.

C. Limitations of the current modeling approach

This study proposed an auditory signal processing model that accounts both for within-channel and across-channel processing in CMR. However, only three basic experiments were considered in order to evaluate the model and to discuss the main principles of auditory processing underlying CMR—in the framework of the model. A number of experimental conditions have been investigated in previous CMR studies, which have not been considered directly in the present study. These studies investigated in much more detail effects of signal frequency (e.g., Schooneveldt and Moore, 1987), masker spectral width (Haggard *et al.*, 1990; Hall and Grose, 1990) and masker spectral level (Moore and Shailer, 1991; Bacon *et al.*, 1997; Cohen, 1991; Hall, 1986; McFadden, 1986), the influence of the envelope statistic of the masker modulator (e.g., Eddins and Wright, 1994; Grose and Hall, 1989; Moore *et al.*, 1990; Hicks and Bacon, 1995), the effect of modulation frequency and modulation depth (Carlyon and Stubbs, 1989; Hall *et al.*, 1996; Lee and Bacon, 1997; Bacon *et al.*, 1997; Verhey *et al.*, 1999; Eddins, 2001), effects of flanking band number and flanking band level (e.g., Hatch *et al.*, 1995; Schooneveldt and Moore, 1987) and other effects. The current version of the model does not include a nonlinear peripheral filtering stage and therefore cannot account for level-dependent cochlear compression and effects associated with it such as level-dependent frequency tuning and suppression. While suppression does not seem to play a role in CMR with the level combinations in the present study (Hall *et al.*, 1984b; Hchooneveldt and Moore, 1987), effects of frequency selectivity certainly do, as was also shown in the present study. However, while corresponding modifications will change the details of the modeling outcomes, the main principles and implications discussed in the present study are expected to remain valid.

A further potential generalization of the model would be to include effects of dichotic presentation of flanker bands on CMR. The size of across-ear effects on CMR (2–3 dB) typically corresponds to that found in monaural across-channel CMR with one flanking band. The idea would be to apply the "central" EC mechanism to the stimuli after consideration of the inputs coming from the two ears. A binaural signal processing model based on the model by Breebaart *et al.* (2001a, b, c) but including a modulation filterbank stage is currently under development.

VIII. SUMMARY AND CONCLUSIONS

- A monaural auditory processing model was proposed that accounts for comodulation masking release (CMR) obtained in perceptual listening tests. The model distinguishes between contributions to CMR from within-channel processing and those resulting from explicit across-channel processing. For the across-channel process, an equalization-cancellation stage was assumed, conceptually motivated by models on binaural processing.
- The model accounts for the main findings in three critical experiments of CMR: (i) CMR with widely spaced flanking bands (where only across-channel processing contributes), (ii) CMR with one flanking band varying in frequency (where within-channel processing dominates at small separations while across-channel processing takes over at large separations), and (iii) CMR obtained in the classical band-widening experiment (where within-channel processing can never be eliminated).
- The simulation results support the earlier hypothesis that (at least) two different processes can contribute to CMR. The within-channel contributions can be as much as 15 dB and is caused by changes of the envelope statistics of the stimulus due to the addition of the signal to the (comodulated) masker—at the output of the auditory filter tuned to the signal frequency. The across-channel process is robust but small (about 2–4 dB) and only observable at small flanker bandwidths (below about 50 Hz).
- Specifically, in the classical band-widening experiment, which originally was used to define CMR as an across-channel process, the simulation results suggest that across-channel processing is not effective, not even at the largest noise bandwidth considered (2000 Hz) where several auditory filters are excited. CMR in this type of stimulus paradigm is dominated by within-channel processes.
- The current implementation of the model does not include a nonlinear, level-dependent cochlear filtering stage which limits its applicability in some of the experimental conditions tested in previous CMR studies. The effect of a level-

dependent frequency selectivity was investigated in one of the experiments of the present study using gammachirp instead of gammatone filters. A more complete implementation in the framework of the whole model is currently under investigation. Overall, the proposed model might provide an interesting framework for the analysis of fluctuating sounds in the auditory system.

ACKNOWLEDGMENTS

We thank our colleagues at the Centre for Applied Hearing Research for many interesting discussions, and John Grose and two anonymous reviewers for their constructive and very helpful criticism. This work was supported by the Danish Research Foundation and the Deutsche Forschungsgemeinschaft (DFG; SFB/TRR 31).

Bacon, S. P., Lee, J., Peterson, D. N., and Rainey, D. (**1997**). "Masking by modulated and unmodulated noise: Effects of bandwidth, modulation rate, signal frequency, and masker level," J. Acoust. Soc. Am. **101**, 1600–1610.

Berg, B. G. (**1996**). "On the relation between comodulation masking release and temporal modulation transfer functions," J. Acoust. Soc. Am. **100**, 1013–1023.

Bernstein, L. R., and Trahiotis, C. (**1996**). "On the use of the normalized correlation as an index of interaural envelope correlation," J. Acoust. Soc. Am. **100**, 1754–1763.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. I. Model structure," J. Acoust. Soc. Am. **110**, 1074–1088.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). "Binaural processing model based on contralateral inhibition. I.Model structure," J. Acoust. Soc. Am. **110**, 1074–1088.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001c**). "Binaural processing model based on contralateral inhibition. I. Model structure," J. Acoust. Soc. Am. **110**, 1074–1088.

Bregman, A. S., Liao, C., and Levitan, R. (**1990**). "Auditory grouping based on fundamental frequency and formant peak frequency," Can. J. Psychol. **44**, 400–413.

Buss, E., and Hall, J. W. (**1998**). "The role of auditory filters in comodulation masking release (CMR)," J. Acoust. Soc. Am. **103**, 3561–3566.

Buss, E., Hall, J., and Grose, J. (**1998**). "Change in envelope beats as a possible cue in comodulation masking release (CMR)," J. Acoust. Soc. Am. **103**, 1592–1597.

Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Carlyon, R., and Stubbs, R. (**1989**). "Detecting single-cycle frequency modulation imposed on sinusoidal, harmonic, and inharmonic carriers," J. Acoust. Soc. Am. **85**, 2563–2574.

Cohen, M. F., and Schubert, E. D. (**1987**). "The effect of cross-spectrum correlation on the detectability of a noise band," J. Acoust. Soc. Am. **81**, 721–723.

Cohen, M. (**1991**). "Comodulation masking release over a three octave range," J. Acoust. Soc. Am. **90**, 1381–1384.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the "effective-Bsignal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am. **99**, 3615–3622.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2893–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," J. Acoust. Soc. Am. **102**, 2906–2919.

Dau, T., Verhey, J., and Kohlrausch, A. (**1999**). "Intrinsic envelope fluctuations and modulation-detection thresholds for narrowband noise carriers," J. Acoust. Soc. Am. **106**, 2752–2760.

Dau, T., Ewert, S. D., and Oxenham, A. J. (**2005**). "Effects of concurrent and sequential streaming in comodulation masking release," in *Auditory Signal Processing -Physiology, Psychoacoustics, and Models*, edited by D. Pressnitzer, A. de Cheveigne, S. MacAdams, and L. Collet (Springer, New York), 335–343.

Derleth, R. P., and Dau, T. (**2000**). "On the role of envelope fluctuation processing in spectral masking," J. Acoust. Soc. Am. **108**, 285–296.

Derleth, R. P., Dau, T., and Kollmeier, B. (**2001**). "Modeling temporal and compressive properties of the normal and impaired auditory system," Hear. Res. **159**, 132–149.

Domnitz, R. H., and Colburn, H. S. (**1977**). "Lateral position and interaural discrimination," J. Acoust. Soc. Am. **61**, 1586–1598.

Durlach, N. (**1960**). "Note on the equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am. **32**, 1075–1076.

Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. **35**, 1206–1218.

Eddins, D. A., and Wright, B. A. (**1994**). "Comodulation masking release for single and multiple rates of envelope fluctuation," J. Acoust. Soc. Am. **96**, 3432–3442.

Eddins, D. A. (**2001**). "Measurement of auditory temporal processing using modified masking period patterns," J. Acoust. Soc. Am. **109**, 1550–1558.

Ernst, S. M. A., and Verhey, J. L. (**2005**). "Comodulation masking release over a three octave range," J. Acoust. Soc. Am. **91**, 998–1006.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

Ewert, S. D., and Dau, T. (**2004**). "Internal and external limitations in amplitude-modulation processing," J. Acoust. Soc. Am. **116**, 478–490.

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am. **112**, 2921–2931.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–65.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Green, D. M. (**1992**). "On the similarity of two theories of comodulation masking release," J. Acoust. Soc. Am. **91**, 1769.

Grose, J. H., and Hall, J. W. (**1989**). "Comodulation masking release using SAM tonal complex maskers: Effects of modulation depth and signal position," J. Acoust. Soc. Am. **85**, 1276–1284.

Grose, J. H., and Hall, J. W. (**1993**). "Comodulation masking release: Is comodulation sufficient?," J. Acoust. Soc. Am. **93**, 2896–2902.

Haggard, M. P., Hall, J. W., and Grose, J. H. (**1990**). "Comodulation masking release as a function of bandwidth and test frequency," J. Acoust. Soc. Am. **88**, 113–118.

Hall, J. W., and Grose, J. H. (**1990**). "Comodulation masking release and auditory grouping," J. Acoust. Soc. Am. **88**, 119–125.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984a**). "Detection in noise by spectro-temporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984b**). "Detection in noise by spectro-temporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.

Hall, J. W., Grose, J. H., and Hatch, D. R. (**1996**). "Effects of masker gating for signal detection in unmodulated and modulated bandlimited noise," J. Acoust. Soc. Am. **100**, 2365–2372.

Hall, J. W. (**1986**). "The effect of across-frequency differences in masking level on spectro-temporal pattern analysis," J. Acoust. Soc. Am. **79**, 781–787.

Hatch, D., Arne, B., and Hall, J. (**1995**). "Comodulation masking release (CMR): Effects of gating as a function of number of flanking bands and masker bandwidth," J. Acoust. Soc. Am. **97**, 3768–3774.

Hicks, M. L., and Bacon, S. P. (**1995**). "Some factors influencing comodulation masking release and across-channel masking," J. Acoust. Soc. Am. **98**, 2504–2514.

Irino, T., and Patterson, R. D. (**1997**). "A time-domain, level-dependent auditory filter: The gammachirp," J. Acoust. Soc. Am. **101**, 412–419.

Langemann, U., and Klump, G. M. (**2001**). "Signal detection in amplitude-modulated maskers. I. Behavioural auditory thresholds in a songbird," Eur. J. Neurosci. **13**, 1025–1032.

Lawson, J. L., and Uhlenbeck, G. E., editors (**1950**). *Threshold Signals, Vol. 24 of Radiation Laboratories Series* (McGraw–Hill, New York).

Lee, J., and Bacon, S. P. (**1997**). "Amplitude modulation depth discrimination of a sinusoidal carrier: Effect of stimulus duration," J. Acoust. Soc. Am. **101**, 3688–3693.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A., and Patterson, R. D. (**1984**). "On the growth of masking asymmetry with stimulus intensity," J. Acoust. Soc. Am. **76**, 739–745.

McFadden, D. (**1986**). "Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band," J. Acoust. Soc. Am. **80**, 1658–1667.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

Meddis, R., Delahaye, R., O'Mard, L., Summer, C., Fantini, D. A., Winter, I., and Pressnitzer, D. (**2002**). "A model of signal processing in the cochlear nucleus: Comodulation masking release," Acta Acust. (Stuttgart) **88**, 387–398.

Moore, B. C. J., and Emmerich, D. S. (**1990**). "Monaural envelope correlation perception, revisited: Effects of bandwidth, frequency separation, duration, and relative level of the noise bands," J. Acoust. Soc. Am. **87**, 2628–2633.

Moore, B. C. J., and Glasberg, B. R. (**1987**). "Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns," Hear. Res. **28**, 209–225.

Moore, B. C. J., and Shailer, M. J. (**1991**). "Comodulation masking release as a function of level," J. Acoust. Soc. Am. **90**, 829–835.

Moore, B. C. J., Hall, J. W., Grose, J. H., and Schooneveldt, G. P. (**1990**). "Some factors affecting the magnitude of comodulation masking release," J. Acoust. Soc. Am. **88**, 1694–1702.

Mott, J. B., McDonald, L. P., and Sinex, D. G. (**1990**). "Neural correlates of psychophysical release from masking," J. Acoust. Soc. Am. **88**, 2682–2691.

Nelken, I., Rotman, Y., and Yosef, O. B. (**1999**). "Responses of auditory-cortex neurons to structural features of natural sounds," Nature (London) **397**, 154–157.

Nieder, A., and Klump, G. M. (**2001**). "Signal detection in amplitude-modulated maskers. II. Processing in the songbird's auditory forebrain," Eur. J. Neurosci. **13**, 1033–1044.

Patterson, R. D., and Moore, B. C. J. (**1986**). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by Moore B. C. J. (Academic, New York), 123–178.

Patterson, R. D., Nimmo-Smith, J., Holdsworth, J., and Rice, P. (**1987**). "An efficient auditory filterbank based on the gammatone function," Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE.

Pressnitzer, D., Meddis, R., Delahaye, R., and Winter, I. M. (**2001**). "Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus," J. Neurosci. **21**, 6377–6386.

Püschel, D. (**1988**). "Prinzipien der zeitlichen Analyse beim Hören (Principles of Temporal Processing in Hearing)," doctoral thesis, Universität Göttingen.

Richards, V. M. (**1987**). "Monaural envelope correlation perception," J. Acoust. Soc. Am. **82**, 1621–1630.

Rotman, Y., Bar-Yosef, O., and Nelken, I. (**2001**). "Relating cluster and population responses to natural sounds and tonal stimuli in cat primary auditory cortex," Hear. Res. **152**, 110–127.

Schooneveldt, G. P., and Moore, B. C. J. (**1987**). "Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band," J. Acoust. Soc. Am. **82**, 1944–1956.

van de Par, S., and Kohlrausch, A. (**1998a**). "Comparison of monaural (CMR) and binaural (BMLD) masking release," J. Acoust. Soc. Am. **103**, 1573–1579.

van de Par, S., and Kohlrausch, A. (**1998b**). "Diotic and dichotic detection using multiplied-noise maskers," J. Acoust. Soc. Am. **103**, 2100–2110.

van de Par, S. (**1998**). "A comparison of binaural detection at low and high frequencies," doctoral thesis, Eindhoven University of Technology.

Verhey, J. L., Dau, T., and Kollmeier, B. (**1999**). "Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model," J. Acoust. Soc. Am. **106**, 2733–2745.

Verhey, J. L., Pressnitzer, D., and Winter, I. M. (**2003**). "The psychophysics and physiology of comodulation masking release," Exp. Brain Res. **153**, 405–417.

Verhey, J. L. (**2002**). "Modeling the influence of inherent amplitude fluctuation simultaneous masking experiments," J. Acoust. Soc. Am. **111**, 1018–1025.

Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.